Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

RECONSTRUCTION IN MENTAL TESTS

AS to the existence of a goodly amount of unrest among psychologists and others who are working in the field of mental tests, there can be no doubt. Such unrest is natural following a period of inflation and is probably symptomatic of a tendency to get back to normal. It is, however, of some importance to watch carefully the process of readjustment, if for no other reason than that it may be as painless as possible and productive of no new disorders that are preventable.

A recent article,[1] clearly a product of this unrest, seems to be traveling over dangerous ground toward a goal that is hardly wholesome. The opinions expressed in this article are shared by enough others in and adjacent to the test field so that its contents assume an importance quite out of proportion to that of the article itself.

The general point of view which is taken is that mental testing is not a descriptive but a technical science; and that since statistical methods as applied to tests have been largely methods borrowed from the descriptive sciences, the time has come to look for fundamental revision in statistical approach. It is urged that every method must aim at the most direct and empirical solution of a problem—"no hypotheses, as thoroughly empirical treatment as may be!" Among other specific assertions, we find, "The actual distribution of various traits is a matter of academic interest only." It is contended that such a revision of methods, aimed at the practical solution of a practical problem, is necessary for the clarification of the total situation.

Before attempting to discuss the truly vital issues involved in such a point of view as this, it may be well to clear the ground by discussing briefly the concrete illustrations in the article. There it is pointed out that a school superintendent who is called upon to make a decision as to the use of this, that, or any test scale, and who is given data as to the "reliability" and "validity" of these scales and information as to the statistical procedure in the construction of this scale, will not be given the information that is really essential to him in making his decision. Of course, this is obvious. But, surprisingly, the conclusion drawn is that the study of the "reliability" and "validity" is futile, and that the construction of test series or tests that approximate as nearly as can be determined true scales of measurement is irrelevant to the problem. The criticism more properly belongs to the advisor of the school superintendent who suggested these three criteria by which to judge the practical usefulness of his tests. The criticism of statistical methods on such grounds is much like finding fault with the esthesiometric index because it had been suggested as a measure of pitch discrimination; or of current quotations in sterling because of their unsuitability as an indication of the selling price of wheat.

---

[1] Pressey, S. L., "Suggestions," *etc. Psychological Review*, Vol. 27, No. 6, page 466.

It is worthy of note, however, that a writer in this field believes that a school superintendent might be thus badly advised. If I may be permitted to depart for a moment from the main discussion of this paper, the fact that such bad advice is obtainable is a matter which psychologists can not much longer ignore. It is the result, to my mind, of two conditions. First, inadequate instruction in statistical technique for students who are doing their work in mental tests; and second, careless editing of psychological and educational periodicals. Not only are most instructional forces rather poorly equipped from the standpoint of personnel to give adequate statistical instruction, but the libraries in many cases lack periodicals and books which are essential for the student in mental tests. To be concrete, Biometrika should be available for first hand reference by such students, and certainly by research workers.

The journals are perhaps even more responsible for the existence of questionable statistical opinion. Articles are all too common which, from a statistical point of view, are distinctly amateurish. One quotation will be illustrative. In a recent article, the subjects which were examined are being described: "In the first group, there were seven boys and eight girls. Of these, two were colored. In the second group, termed the abnormal group, there were nineteen cases. The chronological age ranged from eight to fifteen years. There were ten boys and nine girls. Of these, seven were colored children. *Thus, the two groups represented a chance selection.*" Fifteen pages of the journal were consumed in reporting an experiment based upon such a selection of subjects.

Another type of article which frequently appears is that which naïvely expounds as new some bit of statistical technique which has long had a recognized and reputable standing—without ever a word as to its origin. Thus, we have an article telling how to weight the individual parts of a test to gain maximum diagnostic value without any mention of Spearman's work which was published in 1913. We find articles telling us how to compute a correlation without the use of products, without any acknowledgment of Pearson's or Harris's notes appearing prior to 1910. The demand is made for methods which will show how unmistakably a test scale will set off the lower 15 per cent. or so in scholastic ability; but this technique was devised by Pearson in 1909, its applicability to experimental psychology pointed out by Brown in 1911, and its suitability as an index of diagnostic value in the test field has been discussed by later writers. The authors of such papers themselves can not be held blameless; but after all it is the journals that have brought into existence the reputations to which school superintendents turn for advice.

To come back to the main subject under discussion, the desire for revising our statistical technique to conform with mental testing, not as a descriptive, but as a technical science. There are at least three good reasons why such an attitude is to be deplored.

In the first place, it leads to the giving of approval to various statistical tricks which may be expedient enough in producing results in an immediate and practical situation, but which are actually bad method. Pressey comments on the statistical methods used in the development of the army trade tests as being a step in the right direction. As a matter of fact, from the point of view of method, the trade tests' statistical technique is open to serious criticism. To be sure, the present writer must be held responsible for the statistical method which was used in constructing trade tests, but he would justify it on the ground of practical necessity. Mere practical usefulness does not appear to justify a place for this particular technique as reputable scientific method.

In an earlier article in this JOURNAL,[2] it has been pointed out that the barrenness of the field of mental tests is due in some measure to the extent to which statistical technique has been permitted to compromise the results of test measurement as descriptive material. Certainly, the creation of a "statistics" which shall be simply a tool in the solution of a technical and practical problem can not do other than accomplish further distortion of the actual phenomena which are observed. Pressey would even use his statistics to produce bi-modal distributions in the observed phenomena, because, as he says, "A normal distribution of scores is not to be desired if the greatest efficiency is sought."

To sum up my first point, it seems to me that if we make practical efficiency the criterion by which our statistical technique is judged, we shall encourage the use of innumerable methods, faulty or merely expedient, which have, to be sure, served a purpose, but which can not be genuinely productive in a scientific sense.

---

[2] Vol. XVII, No. 3, p. 57.

In the second place, this point of view results naturally in a tendency to ignore the necessity for analysis, and for the isolation of variable factors in connection with a test problem. Concretely, this tendency finds its expression in the existence of "omnibus" tests and the like, methods which are concerned only with the total score and the immediate desire for successful prediction.

These first two points of criticism gain added significance as they are related to a third point, namely, that this point of view, with its approval of the expeditious, with its tendency to warp the observed data in its statistical machinery, with its inclination to discard as irrelevant and academic the laborious isolation of factors and definition of concepts, must invariably limit the possibility of significant contributions to psychological science from the mental test field.

The future of mental tests, even as applied science, hinges on the capacity of the field to produce contributions that will give us more light on the general problems of mental adjustment. These contributions will be in the highest sense of theoretical importance. It is therefore, to my mind, unfortunate that there should exist a point of view towards statistical technique that will ultimately bring about increased scientific sterility of the field.

The best progress in the field of mental tests can not come from considering that "mental testing is not a descriptive but a technical science." Like any other science, the science of mental tests must be descriptive before it can be technical. We are urged to accept as a slogan: "No hypotheses, as thoroughly empirical treatment as may be!" It would be more suitable to urge at the present time, "No problem that does not test out an hypothesis."

Psychologists could well afford to ignore for the time being the practical applications of mental tests. These have been adequately established and there will always be technicians to work out refinements in practical procedure. The need to-day is for a clarification of the concepts and hypotheses underlying the mental test field; which may in time lead to the development of a theory of measurement that will be consistent with our best knowledge of mental life.

That there must be a change in mental test theory and mental tests methods is evident. The restlessness of workers in the field is certain indication that some change will come. But the change must be an orderly one, based upon scientific principles and upon the most exact methods that exist or that can be devised. "Direct action" can not hold more promise in the realm of science than in the field of politics; and the most productive reconstruction in mental test theory and technique will be one that serves to increase the contribution of mental testing as descriptive science.

BEARDSLEY RUML.

CARNEGIE CORPORATION.

THE TWENTY-NINTH ANNUAL MEETING OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

THE twenty-ninth annual meeting of the American Psychological Association marked an unusually well attended and enthusiastic gathering of American students of psychology. Because of Chicago's central location and probably because of the fact that the American Association for the Advancement of Science also held its meetings in the midwest metropolis, the registration was unusually large. It would be rash of course to attempt to correlate the fresh and eager attitude of the assembled psychologists with the proverbial breezes of Chicago, but possibly there is some relation between the enthusiasm of the psychologists and the remedial effects of time upon the disturbances of war. The calmer consideration of problems scientific and pedagogic gave ample evidence of our passing on from war time events. In truth, the various discussions during the several sessions manifested stronger currents of psychological and general scientific interest than has been the case in the past few years.

The session began on Tuesday morning, December 28, and ended the Thursday morning following. At times the meetings were rather bewildering, since, because of the combined sessions with sections I and sometimes Q of the American Association for the Advancement of Science, there were as many as four sessions at once. All meetings were held in the Law Building of the University of Chicago. The meetings were announced in the programmes as follows: Tuesday morning a session for general psychology; in the afternoon of the same day two sessions were scheduled, one in experimental psychology paralleling a joint session with sections I and Q of the American Association for the Advancement of Science. Wednesday afternoon there was scheduled the address of Dr. Yerkes, the retiring vice-president of section I which was articulated with the psychological association meeting, advertised as a session for social psychology. Dr. Yerkes spoke on "The Relation of Psychology to Medicine," and advocated the establishment of psychology in the medical curriculum as a basic medical science.

On Wednesday afternoon the session on social psychology was followed by the business meeting. In the evening the annual dinner